

Working group 2:  
**Data sharing & re-using**

Stage 4a report:  
**Deliver: What would make  
the biggest impact?**

Last updated:  
**23 May 2025**

Consultation window:  
**26 April -29 May 2025**

## EXECUTIVE SUMMARY

This Stage 4 report by Working Group 2 (WG2) of the Evidence Synthesis Infrastructure Collaborative (ESIC) outlines five high-priority strategies to advance the sharing and reuse of evidence synthesis data globally. These strategies are the culmination of extensive capability assessments, gap analysis, and multi-stage consultations across various regions, sectors, and stakeholder groups. The core aim is to normalize the practice of "studying a question once and reusing the answers across regions, languages, and sectors", thereby ensuring equity, interoperability, and sustainability within evidence systems.

The prioritized strategies are:

- **Building a Federated Repository of Living Evidence Data:** This strategy proposes an interconnected system of evidence repositories, allowing each repository to maintain control over its data while contributing to broader reusability through shared standards and interoperability practices. Key features include federated nodes for sovereign or restricted data, multilingual access, AI-supported tagging and curation tools, and open APIs for cross-platform data integration. This architecture supports interoperability without requiring centralization.
- **Developing Standardized Record Structures and Interoperable Formats:** This strategy focuses on developing shared record structures for evidence synthesis inputs and outputs and adopting machine-readable data formats. This foundational layer aims to reduce redundancy, improve cross-platform data exchange, and enable automation, benefiting both current and legacy datasets through prospective and retrospective conversion.
- **Establishing Metadata standards to facilitate data identification and discoverability:** This strategy introduces modular metadata schemas and classification standards, including multilingual taxonomies and tagging protocols, to enhance evidence classification, searchability, and linkage across platforms. It emphasizes increasing the visibility of underrepresented outputs, such as non-English literature, LMIC evaluations, and grey literature, to foster greater equity and trust.
- **Setting Open Access Standards for Equitable Sharing:** This strategy aims to establish open access policies and standards that ensure equitable sharing and reuse of evidence synthesis data by addressing legal, financial, and structural barriers. It promotes permissive licensing, ethical reuse, and legal clarity, and supports infrastructure for citation, versioning, and governance integration to recognize evidence as a global public good.
- **Ensuring Quality Assurance and Monitoring Systems:** This strategy embeds a tiered validation system alongside quality assurance (QA) and monitoring and evaluation (M&E) mechanisms within the infrastructure. It includes co-developed criteria for assessing completeness, relevance, reliability, and ethical compliance of shared data, and participatory review processes to ensure data integrity, equity, and system accountability. This strengthens user trust and system performance over time.

In addition to the five key strategies, A crucial cross-cutting strategy, **Sustainable Funding and Incentive Alignment**, supports the long-term viability and inclusive participation of all five strategies. This includes establishing multi-donor pooled funds, aligning investment structures, and creating recognition and reward systems for data producers, reviewers, translators, and platform stewards. It also emphasizes equitable access to funding and support for LMIC institutions and under-recognized contributors, along with behavioural incentive models to encourage cultural shifts towards shared infrastructure stewardship.

This report serves as the foundation for upcoming activity-level costing (Stage 4b) and integration with other working groups (Stage 5), culminating in the June 2025 consensus meeting where the full ESIC proposals will be finalized. Many components of these strategies, particularly those focused on metadata standards, record structures, open access, and QA mechanisms, are realistically implementable within 24–36 months, with early pilots possible within the first year. These strategies offer a credible, coordinated, and scalable pathway to make evidence reuse a global norm

## INTRODUCTION

This Stage 4 report presents the final prioritized strategies developed by Working Group 2 (WG2) of the Evidence Synthesis Infrastructure Collaborative (ESIC), focused on data sharing and reuse. Building on the foundational work from earlier stages—mapping current capabilities (Stage 1), identifying gaps (Stage 2), and proposing potential solutions (Stage 3)—this stage aimed to determine which strategies hold the greatest potential for transformative, equitable, and scalable impact.

Informed by consultation and group deliberation, the WG consolidated a wide range of proposals into five core strategies. These strategies reflect the shared vision of making it normal to do core work of evidence synthesis once and reuse it many times across languages and contexts, across regions, languages, and sectors. Each strategy has been assessed using an impact-effort lens and aligned with the broader objectives of the ESIC planning process, including principles of equity, sustainability and interoperability.

This report synthesizes those prioritized strategies, outlines the criteria used for selection, and prepares the ground for the costing and integration work to follow in Stages 4b and 5.

## APPROACH TO PRIORITIZE STRATEGIES AND SUMMARY OF RESULTS

A survey was conducted among WG members using a Likert-type scale with 5 levels. The survey had 7 sections that matched the groups of strategies that were presented in report #3. WG members were asked to rank each strategy in terms of impact and effort, and to provide a rationale for their scores. An impact-effort matrix was produced using the survey results with an average score of 3.5 as the cut-off point between high and low impact or effort (see Annex 2). The strategies were then ranked by average scores (from high to low) and were prioritized in batches:

- **Batch 1 (broad consensus):** All strategies that had a lower limit of an interquartile range for impact of 4 or higher (i.e., 75% of the votes considered the strategy as large or extreme impact) were included.
- **Batch 2 (moderate to high impact and low effort):** All strategies that had an average impact greater than 3.5 and an average effort below 3.5 were also included.

The included strategies were then grouped based on overlaps and complementarity. Incentives to comply with existing infrastructure was prioritized, based on Working Group discussions that identified incentives and behaviour change as critical enablers for uptake. While much of the report focuses on technical infrastructure, the Working Group emphasized that sustained use and alignment with common standards will require intentional attention to cultural, institutional, and motivational incentives.

Finally, the Working Group had a structured discussion to reach consensus on the inclusion of strategies and grouping decisions. Some strategies that were strongly linked and aligned to prioritised strategies were added through this discussion. The final grouped strategies are presented below. Annex 3 outlines the relationship between these strategy groups to the strategies included in stage 3 report. Annex 4 lists strategies that were not prioritized. Note that strategies that were not prioritized does not necessarily mean they were excluded, as many strategies have substantial overlap with prioritized strategies included in this report.

## PRIORITIZED STRATEGIES

Five strategies were finally prioritized and selected:

- Building a Federated Repository of Living Evidence Data
- Developing Standardized Record Structures and Interoperable Formats
- Establishing Metadata standards to facilitate data identification and discoverability
- Setting Open Access Standards for Equitable Data Sharing and Reuse
- Ensuring Quality and Monitoring of Data Sharing and Reuse Systems

The Working Group also identified an important cross-cutting strategy—Sustaining the Infrastructure: Funding Models and Incentives—as essential to enable and support the long-term success of all five strategies.

<b>Strategy 1: Building a Federated Repository of Living Evidence Data</b> <i>Evidence synthesis infrastructure (tools, platforms)</i>	
<b>Description</b>	<p>This strategy proposes a connected system of evidence repositories that work together, rather than a single centralized platform. Each repository keeps control over its own data but contributes to broader efforts to maximize data reusability by aligning, where feasible, with shared standards and interoperability practices. This may include services that help users discover data across repositories, such as optional aggregation of basic metadata but other approaches to facilitating discovery may also be used. The goal is to make structured evidence easier to find, use, and build on, while respecting different data policies and needs.</p>
<b>Problem</b>	<p>Despite the growing volume of evidence, its effective use is hindered by systemic fragmentation and access limitations.</p> <ul style="list-style-type: none"> <li>Existing evidence repositories operate in silos, limiting cross-platform coordination.</li> <li>Many repositories restrict access—sometimes due to legal or institutional barriers that prevent data sharing across borders or in specific formats.</li> <li>There is widespread duplication of effort in data extraction and synthesis stages.</li> <li>The absence of a connected system prevents coordination and discovery across repositories.</li> </ul>
<b>Strategic value</b>	<p>The repository addresses inefficiencies in evidence extraction, duplication of effort, and limited discoverability. Its design promotes legitimacy by embedding transparent governance, adopting open standards, and ensuring all AI-enabled components are tested and validated before scale-up. More than a technical solution, the repository serves as connective infrastructure enabling interoperability across evidence producers, synthesis teams, and policy actors.</p>
<b>Innovation</b>	<p>Unlike isolated platforms or centralized databases, this approach supports a federated system of repositories that can interconnect while maintaining local control. It allows repositories to interconnect without sacrificing control or sovereignty. The system is enhanced through AI-supported tagging, multilingual access, and user-friendly interfaces to make participation accessible across technical capacities.</p>
<b>Equity considerations</b>	<p>This strategy supports equitable participation by allowing local, regional, or sector-specific repositories to operate according to their own data policies while still contributing to a broader ecosystem. The governance approach will aim to support shared leadership, enabling LMICs and community-based repositories to co-lead and contribute to the development of standards. ensures that LMICs and community-based repositories can co-lead and define standards. Multilingual access and open APIs help reduce exclusion based on language, infrastructure, or jurisdiction.</p>
<b>Expected outcomes</b>	<ul style="list-style-type: none"> <li>Faster and more targeted evidence discovery for systematic reviewers, evidence platforms, and policy users, enabled by structured search across interoperable repositories instead of isolated databases.</li> <li>Reduced duplication of synthesis work for evidence producers, researchers, and technical teams, as shared infrastructure makes existing data more visible and reusable across regions and platforms.</li> <li>More inclusive access to underrepresented evidence for LMIC institutions, multilingual users, and regional decision-makers, supported through multilingual and regionally governed nodes.</li> <li>Lower infrastructure and maintenance costs over time for funders, platform hosts, and public institutions, by aligning systems around shared standards, APIs, and governance models.</li> </ul>

### **Description of core components:**

Convening a federated repository of living data involves building an interconnected system of repositories—each maintaining its own governance and structure, while participating in shared standards and infrastructure. The system will include discovery functions to help navigate and access evidence across repositories. The overall system will support interoperability, equity, and local control, while enabling shared access.

At least six core components will be part of this initiative:

- Mapped and Curated Repository Network:** A living map of evidence data repositories, organized by sector, coverage, evidence type, and question type. This includes identifying and curating existing repositories to surface gaps, reduce duplication, and support future linkages.

- **Federated Nodes for Sovereign or Restricted Data:** Interoperable regional or sectoral repositories for cases where data cannot be exported or centrally stored due to policy, legal, or ethical considerations. These nodes are linked to the central system while respecting local ownership.
- **Open API for Cross-Platform Data Integration:** A standardized, open-access application programming interface (API) to enable consistent data exchange across existing and emerging platforms, facilitating real-time integration and interoperability.
- **AI-Enabled Infrastructure for Data Extraction and Curation:** Artificial intelligence tools to support the extraction, structuring, and curation of data entering the repository. These tools assist in preparing content for reuse and improving navigability across repositories.
- **User Interface for Data Exploration and Analysis:** An accessible interface that allows users—especially non-technical audiences—to explore and analyze available data using both quantitative and qualitative tools.
- **Multilingual Access and Navigation:** All repository functionalities will be accessible in multiple languages to reduce access barriers and ensure inclusive usability. This includes multilingual interfaces for search, filtering, and data analysis.

### Key Activities and Sequencing:

This strategy will be implemented through a phased approach.

- **Phase 1: Foundation & Prototyping** will include mapping existing repositories, establishing governance and technical working groups, compiling and validating existing Open APIs and developing additional APIs as needed, and compiling, validating, and adapting AI tagging tools and reuse interfaces. This initial phase will also involve identifying pilot LMIC institutions for early participation.
- **Phase 2: Expansion & Feature Enrichment** will focus on launching 2–3 federated nodes and enhancing existing repository connections. Additional activities will include expanding AI tools with feedback loops, building multilingual interfaces and initiating translation workflows, and developing user-driven data analysis tools with flexible export options for users to apply their own tools. This phase will also involve starting structured training via WG5.
- **Phase 3: Maturity & Integration** will aim to achieve full technical interoperability across nodes, institutionalize governance, metrics, and update cycles, operationalize translation infrastructure and workflows, integrate the repository with global platforms and policymaking tools, and launch a long-term funding and sustainability strategy.

The detailed activities, resource needs, and sequencing for each strategy are outlined in Appendix 2, providing a practical roadmap for phased implementation and costing.

#### Clarifying the Distinctions Between Strategies 2, 3, and 4

While the following three strategies—on record structures, metadata, and open access standards—are closely interrelated, each addresses a distinct layer in the infrastructure required for scalable, trustworthy, and equitable data reuse:

- Strategy 2 (Record Structures and Formats) focuses on the technical structuring of evidence data—how information is formatted, encoded, and transformed for reuse across systems.
- Strategy 3 (Metadata Standards) emphasizes the semantic and discovery layer, enabling classification, filtering, and linkages across repositories using multilingual taxonomies, identifiers, and tagging tools.
- Strategy 4 (Open Access Standards) centres on the legal and policy frameworks—such as licensing, attribution, and permissions—that enable equitable sharing and reuse.

Together, these strategies support interoperability across platforms and across the lifecycle of evidence—from structured extraction to discoverable metadata to open, reusable content. Where appropriate, ESIC will adopt or adapt external standards (e.g., FHIR, RDF, OAI-PMH, AGROVOC) and create specific profiles tailored to the needs of evidence synthesis. Overlaps are intentional to ensure integrated delivery, but implementation efforts will assign clear functional scopes to avoid redundancy.

<b>Strategy 2: Developing Standardized Record Structures and Interoperable Evidence synthesis infrastructure (tools, platforms)</b>	
<b>Description</b>	<p>This strategy addresses the structural fragmentation that limits the efficient use and reuse of evidence synthesis data. It focuses on developing shared record structures for inputs and outputs of syntheses, adopting machine-readable data formats, and enabling both prospective and retrospective conversion of data into these formats. It aims to define common structures for how data is organized (record structures) and how it is shared across platforms (data formats), emphasizing both standardization—ensuring consistency—and interoperability—ensuring compatibility across systems, tools, and sectors. Together, these components provide a foundation for true interoperability across platforms, organizations, and sectors.</p>
<b>Problem</b>	<p>Inadequate interoperability across systems involved in the production and uses of evidence syntheses due to:</p> <ul style="list-style-type: none"> <li>• Lack of consistent / standardized record structures for different inputs and outputs of evidence syntheses.</li> <li>• Incompatible or informal data formats, preventing reuse and platform integration</li> <li>• Vast amounts of legacy data locked in non-standard or narrative-only formats, inaccessible for automated processing.</li> </ul>
<b>Strategic value</b>	<p>Implementing this strategy offers substantial efficiency gains by reducing the need for manual reformatting and enabling the consistent reuse of evidence across diverse tools, platforms, and user groups. A standardized set of record structures and formats lays the foundation for interoperability and ensures that data captured today can be reliably integrated, analyzed, and reused in the future. These standards also serve as a critical enabler for AI-driven tagging, summarization, and automation in evidence synthesis workflows—allowing tools to function more accurately and at scale.</p>
<b>Innovation</b>	<p>This strategy moves beyond conventional narrative reporting guidelines to define structured metadata models that are machine-readable, interoperable, and adaptable across sectors. While there are certain reporting standards across some forms of evidence (e.g., CONSORT, PRISMA), and initiatives that have aimed to achieve interoperability across evidence ecosystems (e.g., EBM on FHIR), there is still plenty of room to improve interoperability across all systems, sectors and forms of evidence, as well as all inputs and outputs of evidence syntheses. Developing standard record structures and data formats will be a unique strategy to create a foundational infrastructure layer for a connected, future-ready evidence ecosystem.</p>
<b>Equity considerations</b>	<p>Standardized formats lower the technical barriers to entry, enabling LMIC institutions and smaller organizations to contribute and reuse data using open tools and templates. Retrospective conversion efforts ensure that older data—often produced in low-resource settings—is not excluded from the shift toward structured and automated synthesis workflows. Supporting multilingual documentation and community co-design will further strengthen accessibility and equity.</p>
<b>Expected outcomes</b>	<ul style="list-style-type: none"> <li>• Faster integration of evidence across tools and systems for synthesis platforms, data curators, and automation developers, reducing manual formatting and enabling smoother workflows.</li> <li>• Lower technical barriers for participation for LMIC institutions, smaller research teams, and regional evidence networks, through the use of adaptable and standardized formats.</li> <li>• Improved data provenance and trust for evidence users, policymakers, and reviewers, with structured records supporting version tracking and transparency in how evidence is curated and shared.</li> <li>• Greater consistency in how synthesis inputs and outputs are stored, exchanged, and reused by platform developers, funders, and cross-sector partners, enhancing long-term interoperability</li> </ul>

### Description of core components:

At least four core components will be part of this initiative:

- A standardized record structure for evidence synthesis inputs (e.g., primary studies, evaluation reports) and outputs (e.g., synthesis reports, quality appraisals, data extracted, etc.).
- Adopting an interoperable machine-readable data format that would enable free exchange across platforms.

- Systems that allow prospective conversion of data used and produced by evidence syntheses into the interoperable data format defined.
- Systems that support the retrospective conversion of data into the interoperable data format.

### Key Activities and Sequencing:

This strategy will be implemented through a phased approach.

- In **Phase 1: Foundation & Prototyping**, activities will include establishing technical and advisory groups to oversee standard development, mapping existing standards and tools for evidence data sharing, developing initial record structure standards for inputs and outputs, drafting interoperable data formats and exchange protocols, and prototyping tooling for data conversion and structure validation.
- **Phase 2: Expansion & Feature Enrichment** will focus on building and testing retrospective conversion toolkits. This phase will also involve piloting conversion workflows with selected institutions, such as LMIC partners and sectoral repositories.
- Finally, **Phase 3: Maturity & Integration** will aim to launch open consultation and finalize standards based on feedback. It will also institutionalize standards through integration into platform governance and training programs.

<b>Strategy 3: Establishing Metadata standards to facilitate data identification and discoverability</b> <i>Evidence synthesis infrastructure (tools, platforms)</i>	
<b>Description</b>	<p>This strategy aims to develop and implement metadata standards that will allow for the consistent and efficient identification, classification, and discoverability of data used and produced in evidence syntheses. It supports interoperability across platforms and enhances reuse by enabling structured tagging and classification of studies, reports, and other outputs.</p> <p>This also include the adoption or development of common typologies and taxonomies to classify evidence consistently across domains, including by topic, method, population, and context. These shared classification structures will support multilingual discovery, filtering, and reuse across federated platforms.</p>
<b>Problem</b>	<p>Even when evidence is well-structured, it often remains difficult to find, interpret, or reuse due to missing or inconsistent metadata. The lack of shared identifiers, classification systems, and descriptive tags makes it challenging to connect related studies and synthesis outputs across platforms. Systematic reviewers, in particular, must often rely on time-consuming and error-prone free-text searches due to the absence of structured metadata. This not only reduces precision but increases screening burden and introduces risks of omission. This is especially true for non-English sources, grey literature, and evidence generated in LMICs, which are frequently under-indexed or omitted entirely. Without standardized metadata, users cannot reliably filter by study design, outcome type, population group, or topic area. It also becomes difficult to assess the provenance, version history, or quality status of shared data, limiting trust and usability across the evidence ecosystem.</p>
<b>Strategic value</b>	<p>Metadata is the foundation of discoverability. Implementing shared metadata schemas and classification standards will improve the ability of users to find the evidence they need across platforms, languages, and regions. It enables repositories to index and retrieve relevant evidence more effectively, allows search tools to sort and filter results, and supports AI-driven automation and quality control. Standardized metadata will also promote transparency, by ensuring that users can track the source, version, and curation history of each record.</p> <p>High-quality metadata enables discovery, reuse, and equity in access. However, its creation requires time and expertise, and without guidance, automation, and incentive, may risk becoming a bottleneck for evidence producers. The strategy therefore emphasizes both ease of implementation and shared responsibility across the ecosystem.</p>
<b>Innovation</b>	<p>This strategy introduces a structured, multilingual vocabulary for key dimensions of evidence synthesis, including study type, topic, outcome, and population group. It includes tools to support both manual and AI-assisted tagging, as well as protocols for assigning persistent identifiers (e.g., DOIs) and linking related content. By leveraging upstream standards and automation where possible, the approach reduces redundancy and enables metadata generation during existing workflows. The standards will be designed to function across sectors and geographies, and may be applied to previously structured records where feasible to improve their discoverability.</p>

Equity considerations
Metadata standards will be inclusive by design, with multilingual fields and community input. By increasing the visibility of underrepresented outputs—such as non-English literature, LMIC evaluations, and grey literature—this strategy supports equitable access to evidence. It also lowers technical barriers to participation by providing templates and tagging tools suitable for institutions with limited infrastructure.
Expected outcomes
<ul style="list-style-type: none"> <li>• Faster and more accurate evidence retrieval for systematic reviewers, data scientists, and decision-makers, as structured metadata enables precise filtering by topic, population, method, and outcome.</li> <li>• Improved equity in visibility and reuse for LMIC institutions, multilingual users, and civil society organizations, with multilingual tagging and standard classifications increasing access to grey literature and under-indexed evidence.</li> <li>• Reduced workload and bottlenecks for evidence producers, research teams, and platform developers, by integrating metadata into upstream workflows and supporting automation and reuse across platforms.</li> <li>• Stronger transparency and trust for end-users, funders, and evaluators, as standardized metadata supports tracking of provenance, version history, and quality assurance status.</li> </ul>

### Description of core components:

The following components are essential to ensure consistent, multilingual, and interoperable metadata practices across platforms, enabling effective discovery, filtering, and reuse of evidence:

- **Modular metadata Schema and Classification Standards:** A metadata schema covering study/article IDs, classifications (e.g., design, population, outcome type), thematic tags (e.g., SDGs), and quality fields. The metadata should be modular, allowing to be adapted from different types of records (inputs and outputs of evidence syntheses) and platforms. To reduce bottlenecks for evidence producers, the schema will prioritize essential fields and support incremental enrichment over time. Tagging should be as close to upstream workflows as possible and supported by automation where feasible. Where feasible, a common core module can be established to promote cross-sector interoperability, while allowing additional sector-specific or regional modules to extend the schema based on local needs or evidence types. This balance supports consistency without imposing uniformity.
- **Multilingual and Cross-Sector Alignment of typologies:** Controlled vocabularies and metadata fields translated and validated across multiple languages and sectors (e.g., health, education, environment). This alignment enables consistent classification, filtering, and discovery of evidence across domains, while ensuring accessibility for diverse users. Where appropriate, sector-specific taxonomies should be mapped to shared typologies to support cross-sector interoperability without sacrificing domain relevance.
- **Tools that enable transformation to support taxonomy alignment:** Manual and AI-assisted tools for applying metadata, with built-in QA features to ensure tagging is consistent, traceable, and updateable, and can be translated to multiple available taxonomies. These tools should be designed to minimize burden on producers and facilitate reuse of metadata across platforms, including alignment with external standards (e.g., Dublin Core, Schema.org) to reduce duplication.
- **Transparent Provenance and Versioning Records:** Metadata fields that record changes, updates, and curation steps—allowing users to understand what transformations the data has undergone.
- **Identifier Systems and Linkages:** Guidance and tooling for assigning and maintaining persistent identifiers (e.g., DOIs), with linkages between studies, syntheses, and derivative outputs.
- **Metadata Enrichment of Legacy Records:** Where feasible, tagging standards may be applied to previously structured records to enhance discoverability.

### Key Activities and Sequencing:

This strategy will be implemented through a phased approach.

- **In Phase 1: Foundation & Prototyping,** activities will include establishing a metadata working group and regional advisors, defining the metadata schema and classification system, mapping and reviewing existing metadata standards and typologies, and developing a strategy for mapping legacy metadata into a unified framework. This phase will also involve drafting identifier guidance and linkage protocols.



- **Phase 2: Expansion & Feature Enrichment** will focus on developing tagging tools, including AI-assisted modules. This phase will also involve conducting multilingual validation and testing, and piloting the tagging system across selected ESIC-aligned repositories.
- Finally, **Phase 3: Maturity & Integration** will aim to finalize standards through open consultation and feedback. It will also institutionalize the metadata schema and QA practices across ESIC tools, and maintain the multilingual metadata dictionary and update protocols.

<b>Strategy 4: Setting Open access standards for equitable data sharing and reuse</b> <i>Evidence synthesis infrastructure (tools, platforms)</i>	
<b>Description</b>	<p>This strategy aims to establish open access standards and policies that ensure equitable sharing and reuse of evidence synthesis data across contexts, sectors, and user groups. It addresses legal, financial, and structural barriers to access, promotes clear licensing and attribution frameworks, and supports system-wide norms that recognize data as a global public good.</p>
<b>Problem</b>	<p>Despite growing volumes of evidence synthesis data, access remains restricted due to legal, technical, or cost-related barriers. Without common policies and standards, open access practices remain inconsistent and exclusionary. Key challenges include:</p> <ul style="list-style-type: none"> <li>• Unclear or restrictive licensing terms limiting reuse and redistribution</li> <li>• Paywalls or subscription models creating inequities in data access</li> <li>• Lack of shared norms for attribution, versioning, or derivative use</li> <li>• Fragmented policies across repositories, institutions, and publishers</li> <li>• Limited infrastructure to support equitable open access across languages and sectors</li> </ul>
<b>Strategic value</b>	<p>Open access is a precondition for equity and interoperability. Establishing common policies and licensing frameworks enables data to move freely across platforms, regions, and use cases. It ensures that evidence generated with public funding or intended for public benefit is available to all users—especially those in low-resource settings. Open access also facilitates reuse, integration, and innovation, enabling researchers, policymakers, and communities to build upon shared knowledge.</p>
<b>Innovation</b>	<p>This strategy goes beyond promoting openness in principle by defining concrete licensing options, conditions for responsible reuse, and pathways for integrating open access into platform architecture and governance. It promotes the adoption of permissive licenses (e.g., CC-BY, CC0), while accounting for ethical, legal, and institutional sensitivities. It supports shared norms for citing reused content, respecting attribution, and ensuring integrity across derivative works.</p>
<b>Equity considerations</b>	<p>Open access is essential for leveling the playing field. By removing financial and technical access barriers, this strategy enables participation from LMICs, community-based organizations, and early-career researchers. It ensures that publicly funded evidence does not remain siloed in institutions with exclusive access. It also supports multilingual access, local ownership, and ethical sharing practices, including data sovereignty where appropriate.</p>
<b>Expected outcomes</b>	<ul style="list-style-type: none"> <li>• Broader and more equitable reuse of evidence for policymakers, researchers, and civil society actors, as clear licensing terms reduce legal and technical barriers across regions and sectors.</li> <li>• Greater participation from LMICs and smaller institutions, supported by open access models that lower paywalls and permission barriers for early-career researchers, regional platforms, and local evaluators.</li> <li>• Increased transparency and accountability for evidence users, data producers, and synthesis curators, with standard attribution and licensing improving visibility of evidence origin and downstream use.</li> <li>• Stronger alignment with open science and funder mandates for public institutions, academic organizations, and grantees, enabling compliance while promoting public-good knowledge sharing.</li> </ul>



## Description of core components:

The following components define the building blocks of an equitable open access ecosystem for data sharing and reuse:

- **Open Licensing Standards:** Guidance on recommended licenses (e.g., CC-BY, CC0) and conditions of reuse and text data mining, adapted for synthesis data and evidence workflows.
- **Infrastructure Support for Access:** Platform-level capabilities for downloading, linking, versioning, and translating open datasets across languages and formats, and supporting organizations in making all their products open access.
- **Legal and Ethical Guidance for Reuse:** Templates and policies to ensure compliance with privacy, IP, and data protection laws—particularly when sharing sensitive or context-specific evidence.
- **Open Access Integration into Platform Governance:** Roles and responsibilities for implementing, monitoring, and updating open access practices, embedded in repository and WG-level oversight structures.

## Key Activities and Sequencing:

This strategy will be implemented through a phased approach.

- In **Phase 1: Foundation & Prototyping**, activities will include convening an open access advisory group, drafting open licensing and attribution guidance, and conducting a legal review of reuse models across jurisdictions.
- **Phase 2: Expansion & Feature Enrichment** will focus on building platform capabilities for license display, citation export, and open access tagging. This phase will also involve developing and piloting multilingual access protocols, and hosting consultations with LMIC institutions and data producers on access priorities.
- Finally, **Phase 3: Maturity & Integration** will aim to integrate licensing and attribution norms into platform governance frameworks, monitor and evaluate open access implementation across ESIC systems, and maintain public guidance materials and licensing templates.

Strategy 5: Ensuring Quality and Monitoring of Data Sharing and Reuse Systems	
Evidence synthesis process (methods; training; learning; sharing; convening)	
<b>Description</b>	This strategy aims to embed a tiered validation system alongside quality assurance (QA), monitoring, and evaluation (M&E) mechanisms within the infrastructure and processes that support data sharing and reuse in evidence synthesis. The tiered approach enables different levels of review—ranging from automated checks and methodological audits to participatory peer review—matched to the risk profile, context, and intended reuse of the data. It seeks to ensure that shared data is not only accessible but also trustworthy, up-to-date, ethically managed, and reusable with confidence by evidence producers and users. Recognizing that quality assurance and evaluation processes will never meet all expectations in all contexts, the strategy also emphasizes continuous improvement—establishing mechanisms to assess system performance, equity impacts, and adherence to shared standards, while leaving room to resolve tensions and refine evaluation over time.
<b>Problem</b>	Data sharing systems often lack built-in mechanisms to assess and maintain the quality, relevance, and ethical use of the data they host. Without systematic M&E, users face difficulties in judging whether shared data is current, accurate, or responsibly curated. Key challenges include: <ul style="list-style-type: none"><li>• Lack of common QA processes for evaluating shared data</li><li>• Inconsistent application of quality standards across repositories</li><li>• Absence of feedback loops for reporting and correcting data issues</li><li>• Limited ability to monitor system usage, equity, and performance</li><li>• Unclear accountability for data quality, ethics, and legal compliance</li></ul>
<b>Strategic value</b>	Embedding QA and M&E into infrastructure improves trust, usability, and sustainability. It enables: <ul style="list-style-type: none"><li>• Reliable data reuse by ensuring integrity, transparency, and recency</li><li>• Early detection of quality risks and gaps in system performance</li><li>• Stronger alignment with ethical and legal frameworks</li><li>• Greater stakeholder confidence in shared systems</li><li>• Continuous learning and adaptive improvements</li></ul>

It also supports funders, platforms, and LMIC institutions in documenting value and building accountability.
<b>Innovation</b>
This strategy goes beyond static QA checklists by introducing dynamic feedback mechanisms, periodic system reviews, and stakeholder-led learning loops. It promotes shared responsibility for quality, supports decentralized governance models, and includes participatory tools for surfacing user insights, flagging risks, and guiding improvements. The strategy leverages existing frameworks (e.g., SHOW ME principles) while adapting them for ongoing, system-wide application.
<b>Equity considerations</b>
Robust QA and M&E systems help ensure that data from LMICs is not unfairly flagged or excluded due to differences in format or reporting practices. By developing transparent, locally informed standards, and involving LMIC actors in defining quality benchmarks and M&E indicators, this strategy builds inclusive trust. It also supports monitoring equity in access, participation, and benefit-sharing across data systems.
<b>Expected outcomes</b>
<ul style="list-style-type: none"> <li>• Improved trust and reliability of shared data for decision-makers, synthesis platforms, and evidence users, through transparent validation processes and visible quality indicators.</li> <li>• More efficient reuse of high-quality evidence for reviewers, curators, and guideline developers, as tiered QA systems help users filter by relevance, rigor, and readiness for decision-making.</li> <li>• Faster identification and correction of quality issues for data contributors, repository managers, and QA teams, enabled by integrated feedback loops and user-driven reporting mechanisms.</li> <li>• Better system-wide performance tracking for funders, governance bodies, and coordination platforms, with M&amp;E frameworks generating data on usage, equity, and adherence to shared standards.</li> </ul>

### Description of core components:

The following components ensure that systems for data sharing and reuse remain high-quality, trusted, and responsive to user needs:

- **Tired validation Standards and Guidelines:** Co-developed criteria and templates to assess the completeness, relevance, reliability, and ethical compliance of shared data.
- **Monitoring and Evaluation Framework:** A set of indicators, dashboards, and review mechanisms designed to assess data usage, validation uptake, access equity, and alignment with shared standards. This includes tools to gather structured feedback from users and contributors, mechanisms to surface emerging issues, and processes for governance bodies to act on evaluation findings. Recognizing the diversity of platforms and contexts, the framework will be implemented with flexibility, allowing for local adaptation and periodic refinement.
- **Feedback and Reporting Channels:** Tools to allow users to flag issues with data quality, usage barriers, or policy misalignment—feeding into continuous improvement processes.
- **Participatory Review Processes:** Structured opportunities for interest-holder input on what quality and ethical data use means across sectors and regions.
- **Integration with System Governance:** Clear responsibilities for implementing QA/M&E protocols at both the platform and institutional levels, aligned with WG2-wide governance models.

### Key Activities and Sequencing:

This strategy will be implemented through a phased approach.

- In **Phase 1: Foundation & Prototyping**, activities will include establishing a technical QA and M&E advisory group, developing draft quality assurance standards, and defining indicators for system monitoring and ethical adherence.
- **Phase 2: Expansion & Feature Enrichment** will focus on building user feedback and issue-reporting interfaces. This phase will also involve piloting QA and M&E systems in selected repositories, and conducting participatory review sessions with LMIC and sectoral partners.
- Finally, **Phase 3: Maturity & Integration** will aim to integrate QA/M&E protocols into platform governance and user workflows, maintain and update monitoring dashboards and QA tools, and conduct periodic system-wide quality and equity evaluations.

## CROSS-CUTTING STRATEGY

In addition to the five technical strategies, the report identifies a critical cross-cutting condition: Sustainable Funding and Incentive Alignment. This strategy is foundational to the long-term success, equity, and scalability of the entire infrastructure. It is considered “cross-cutting” because its impact is not limited to a single domain—instead, it underpins the implementation, sustained use, and institutional adoption of all other strategies. Without reliable financing and well-aligned incentives, even the most innovative efforts—whether federated repositories, standardized formats, enriched metadata, or open access systems—are unlikely to achieve widespread uptake, maintenance, or equitable participation, particularly from LMICs and under-recognized contributors.

<b>Sustaining the Infrastructure: Funding Models and Incentives</b> <i>Evidence synthesis infrastructure (tools, platforms)</i>	
<b>Description</b>	<p>This cross-cutting strategy focuses on establishing the long-term financial and motivational systems needed to sustain the infrastructure, processes, and collaborations envisioned in the ESIC planning process. It aims to secure sustainable, multi-source funding and to align incentives across sectors, regions, and stakeholder groups to ensure ongoing participation in data sharing, reuse, and system upkeep.</p> <p>In addition to financial mechanisms, this strategy addresses the behavioral and cultural dimensions of infrastructure use—recognizing that incentives are a key lever for changing norms and encouraging consistent participation across context.</p>
<b>Problem</b>	<p>Despite broad recognition of the value of evidence sharing and synthesis, funding remains fragmented, short-term, and project-based. This undermines the scalability and sustainability of infrastructure investments.</p> <p>Key challenges include:</p> <ul style="list-style-type: none"> <li>• Lack of long-term funding to maintain and evolve core infrastructure</li> <li>• Fragmentation of funding streams, leading to duplication and inefficiencies</li> <li>• Absence of pooled or joint mechanisms to support shared infrastructure</li> <li>• Limited incentives for data contributors, curators, or LMIC institutions</li> <li>• Mismatched reward systems that prioritize publication over sharing and reuse</li> </ul> <p>Even where infrastructure exists, uptake is often hindered by low motivation, institutional inertia, and cultural resistance.</p>
<b>Strategic value</b>	<p>Sustainable infrastructure requires sustainable support. This strategy enables long-term funding, coordination among donors, and clear incentives for institutions and individuals to share, reuse, and enrich data. It supports LMIC leadership, reduces redundancy, and improves returns on investment. By aligning incentives with behaviors like contributing back to repositories, curating structured data, and using common standards, it helps normalize collaboration and long-term system use.</p>
<b>Innovation</b>	<p>This strategy introduces a shared funding architecture, including pooled donor mechanisms, coordinated grant structures, and co-financing models tailored for global public goods. It also defines incentive models that go beyond financial compensation—such as attribution, recognition in funding decisions, integration into research assessment, and performance metrics for repositories and institutions.</p> <p>It also considers upstream behavioral nudges, such as rewarding those who extract standard metadata fields even when not immediately relevant to their own analysis—on the principle that doing the extra 10% enables broader reuse for the collective good.</p>
<b>Equity considerations</b>	<p>Equity is central to both the funding and incentive models. This strategy emphasizes funding models that actively include LMIC partners—not only as beneficiaries, but as co-leaders in governance, implementation, and innovation. Incentive systems will also reflect contributions from curators, translators, local reviewers, and regional institutions that are often under-recognized but essential for system functioning</p>
<b>Expected outcomes</b>	<ul style="list-style-type: none"> <li>• Multi-donor pooled funds and shared grant frameworks supporting infrastructure over time.</li> <li>• Transparent cost models and sustainability plans embedded in each platform.</li> <li>• Reward systems that recognize and incentivize data sharing and reuse.</li> <li>• Equitable access to funding opportunities across geographies and institution types.</li> <li>• Reduced duplication and greater coordination among funders and stakeholders</li> </ul>

## Description of Core Components:

The following components support the long-term sustainability and motivation needed to make the ESIC infrastructure viable and inclusive:

- **Shared Funding Mechanisms:** Multi-donor pooled funds and aligned investment structures to support platform infrastructure, governance, and operations.
- **Sustainability Models and Costing Plans:** Platform-specific business models, including cost-sharing approaches, service models, and public-good justifications.
- **Aligned Incentives for Contributors:** Recognition and reward systems for data producers, reviewers, translators, and platform stewards, integrated into research funding and assessment frameworks.
- **Equitable Access to Funding and Support:** Mechanisms to ensure that LMIC institutions, early-career professionals, and non-traditional actors have equitable access to grants, leadership roles, and capacity investments.
- **Coordination and Accountability Structures:** Mechanisms to ensure funders and stakeholders move in aligned directions, reducing duplication and improving return on investment.
- **Behavioural Incentive Models and Culture Change Tools:** Development of tools and policies that motivate behavioural alignment with shared infrastructure—such as recognition for tagging structured data, contributing to core outcome sets, or prioritizing standard formats even when not required by journals.

## Key Activities and Sequencing:

This cross-cutting strategy will be implemented through a phased approach.

- In **Phase 1: Foundation & Prototyping**, activities will include mapping existing funding flows and gaps across ESIC-aligned infrastructure, convening funders to design pooled or coordinated funding models, and drafting an incentive framework and potential recognition models.
- **Phase 2: Expansion & Feature Enrichment** will focus on piloting shared grant mechanisms across multiple funders. This phase will also involve developing costing templates and sustainability plans for each platform, launching incentive pilots (e.g., citation tracking, recognition in grant calls), and co-designing behavioural incentive pilots (e.g., recognition for data curation, open tagging, or use of core outcome sets) with institutions and funders.
- Finally, **Phase 3: Maturity & Integration** will aim to institutionalize funding and incentive models across ESIC platforms, monitor and evaluate sustainability and equity outcomes, and adapt and scale up funding and incentive models based on feedback.

## CONCLUSION

This Stage 4 report marks a key milestone in the planning process of ESIC Working Group 2. It consolidates and prioritizes a set of five actionable strategies—each focused on strengthening the infrastructure, standards, and systems needed to make data sharing and reuse a global norm. Together, these strategies reflect a coherent, equitable, and forward-looking vision of evidence synthesis infrastructure that is interoperable, inclusive, and capable of supporting timely and trustworthy decision-making. The strategies outlined here have been shaped through iterative consultation, prioritization exercises, and alignment with ESIC-wide objectives. While the vision remains ambitious, the timelines and activities defined for each strategy reflect an optimistic but feasible implementation pathway. Several strategies—particularly those focused on metadata standards, record structures, open access, and QA mechanisms—can be substantially implemented within 24–36 months, with early pilots and institutional integration already possible within the first year. More complex infrastructure elements, such as the federated repository and long-term funding systems, will require longer-term collaboration and support, but their foundational elements are also achievable within the 3-year horizon.

## NEXT STEPS

The next steps will focus on finalizing activity-level costing (Stage 4b) and integrating WG2 strategies with those of other Working Groups as part of a shared roadmap for funders and implementation partners (Stage 5). These will culminate in the June consensus meeting, where the full ESIC proposals will be presented.

## APPENDICES

**Appendix 1.** Roadmap integration table

Problem	Solution			Expected outcome	Success/ destination	Synergies
What is the problem the solution is intended to address or resolve?	Solution name and short description	Type of solution: (a) evidence synthesis infrastructure (b) evidence synthesis process (c) projects	Why is the solution innovative?  (e.g. it's a novel approach that deviates from conventional methods; it presents clear advantages over existing options; it presents a scalable approach that can extend benefits beyond initial scope?	What are the expected changes or benefits to result from implementing the solution? (How does it contribute to cheaper, faster, better synthesis or the SHOW ME principles? Mention the specific type of interest-holders that will be impacted. Link back to the problem statement.	What would success look like if this solution was implemented? Please be as S.M.A.R.T. as possible.	What are the connections to solutions proposed by other WGs  (a) Dependency: (development/implementation of the solution is dependent upon another WG solutions (b) Complementarity: (the solution is helpful or reinforcing in operation. It will have a symbiotic relationship with other solutions
Existing evidence repositories operate in silos and some of them with limited access to interest holders	Building a federated repository of living evidence data  A federated repository that groups a set of comprehensive repositories covering all sectors, forms of evidence and types of question	Evidence synthesis infrastructure (tools, platforms)	While there are multiple repositories of evidence, there is no federated repository that could cover all different sectors, forms of evidence and types of question that facilitate integration and user experience across the existing evidence architecture.	<ul style="list-style-type: none"> <li>• Faster and more targeted evidence discovery for reviewers and policy users.</li> <li>• Reduced duplication of synthesis work across regions and platforms.</li> <li>• Inclusive access for LMICs and multilingual users via regionally governed nodes.</li> <li>• Lower infrastructure costs through aligned APIs and governance models.</li> </ul>	<p>The solution might contribute to a new global infrastructure that would:</p> <ul style="list-style-type: none"> <li>- Make easy the access to the right evidence in one place for decision makers.</li> <li>- Reduce drastically the amount of duplication across evidence producers</li> <li>- Increase coordination across evidence synthesis community.</li> </ul>	<p>Complementarity: The solution includes the uses of AI to conduct curation and data extraction (WG3)</p> <p>Complementarity: The interfaces of the repository should also be functional on how it best serve evidence users (WG1)</p>

Inadequate interoperability across systems involved in the production and uses of evidence syntheses	<p>Developing standardized records structures and data formats to facilitate interoperability</p> <p>Developing shared record structures for inputs and outputs of syntheses and adopting machine-readable data formats.</p>	<p>Evidence synthesis infrastructure (tools, platforms)</p> <p>Evidence synthesis process (methods; training; learning; sharing; convening)</p>	Developing standard record structures and data formats will be a unique strategy to create a foundational infrastructure layer for a connected, future-ready evidence ecosystem.	<p>Faster integration across synthesis tools, reducing manual formatting.</p> <ul style="list-style-type: none"> <li>• Lower technical barriers for LMIC institutions and smaller teams.</li> <li>• Structured records improve provenance and transparency.</li> <li>• Long-term interoperability across platforms and funders.</li> </ul>	The solution should significantly enable interoperability across systems, facilitating the reuses of data	<p>Complementarity: The definition of reporting standards should also have a strong methodological component (WG4).</p> <p>Complementarity: AI tools can facilitate the prospective and retrospective conversion of data into interoperable formats (WG3).</p> <p>Dependency: A strong governance would facilitate the adoption of these standards across different players (GPG)</p>
Lack of shared identifiers, classification systems, and descriptive tags that makes it challenging to connect related studies and synthesis outputs across platforms	Developing metadata standards to facilitate data identification and discoverability	Evidence synthesis infrastructure (tools, platforms)	The standards will be designed to function across sectors and geographies, and may be applied to previously structured records where feasible to improve their discoverability. No similar initiative has been develop to present.	<ul style="list-style-type: none"> <li>• Faster and more precise retrieval for systematic reviewers and policymakers.</li> <li>• Increased visibility of LMIC and non-English evidence.</li> <li>• Reduced producer burden through integrated, automated tagging.</li> <li>• Strengthened transparency via provenance and version tracking.</li> </ul>	Implementing shared metadata schemas and classification standards will improve the ability of users to find the evidence they need across platforms, languages, and regions, as well as facilitating its reuses for different purposes	<p>Complementarity: Agreeing on common typologies/taxonomies might entail a significant amount of methodological expertise (WG4).</p> <p>Dependency: A strong governance would facilitate the adoption of these standards across different players (GPG)</p>

Despite growing volumes of evidence synthesis data, access remains restricted due to legal, technical, or cost-related barriers. Without common policies and standards, open access practices remain inconsistent and exclusionary.	Open access standards for equitable data sharing and reuse	Evidence synthesis infrastructure (tools, platforms)	This strategy goes beyond promoting openness in principle by defining concrete options to promote open science in the context of data sharing and reuse.	<ul style="list-style-type: none"> <li>• Reduced access barriers for LMICs and community organizations.</li> <li>• Increased participation from early-career and regional researchers.</li> <li>• Enhanced visibility and reuse through standard citation/licensing.</li> <li>• Compliance with funder and institutional mandates for open science.</li> </ul>	Open access is a precondition for equity and interoperability. With a strong open-access strategy, we will ensure that evidence generated with public funding or intended for public benefit is available to all users—especially those in low-resource settings.	Dependency: A strong governance would be needed to engage multiple actors in committing to open-access principles (GPG)
Data sharing systems often lack built-in mechanisms to assess and maintain the quality, relevance, and ethical use of the data they host.	Ensuring quality and monitoring and evaluation of data sharing/reusing systems	Evidence synthesis process (methods; training; learning; sharing; convening	This strategy goes beyond static QA checklists by promoting shared responsibility for quality and ensuring trust that would enable data reuse.	<ul style="list-style-type: none"> <li>• Increased user trust through visible validation and quality indicators.</li> <li>• More efficient reuse via tiered quality filtering.</li> <li>• Proactive identification and resolution of quality gaps.</li> <li>• Equity monitoring across use, access, and benefit-sharing.</li> </ul>	Embedding quality assurance mechanisms and M&E strategies into the evidence synthesis infrastructure will significantly increase trust which would encourage the sharing and reuses of data.	<p>Complementarity: Setting up robust quality assurance standards might require important agreement across methods (WG4).</p> <p>Complementarity: Building trust across system players will also be achieved by a strong governance of the system (GPG)</p>



## Appendix 2 : Key Activities and Sequencing:

### Strategy 1: Build a Federated Repository of Living Evidence Data

- Phase 1: Foundation & Prototyping (0–9 months)

Activity	Duration	Resource
Map existing repositories	2 months	People (research / data analysts), Tools (search hardware, repository mapping tools), Capacity development (criteria refinement workshops)
Establish governance and technical working groups	3 months	People (governance coordinators, cross-sector experts)
Compile / validate Existing Open API's and develop additional ones as needed	6 months	People (software developers, data architects), Tools (API frameworks, cloud storage platforms)
Compile, validate, and adapt AI tagging tools and reuse interfaces	6 months	People (AI specialists, UX designers), Tools (machine learning subscriptions, prototype tools)
Identify pilot LMIC institutions	2 months	People (regional engagement leads), Capacity development (introductory onboarding and guidance documents)

- Phase 2: Expansion & Feature Enrichment (9–24 months)

Activity	Duration	Resource Type
Launch 2–3 federated nodes and enhanced repository connections	6 months	People (regional data infrastructure teams), Tools (federation middleware, data sync tools)
Expand AI tools with feedback loops	6 months	People (AI engineers, QA reviewers), Tools (model monitoring systems, tagging engines)
Build multilingual interfaces and initiate translation workflows	6 months	People (localization experts, UX researchers), Tools (translation platforms, multilingual interface software)
Build interface for user-driven data analysis, with flexible export options for users to apply their own tools	6 months	People (frontend/backend developers), Tools (data visualization libraries, analysis dashboards)
Start structured training via WG5	3 months	Capacity development (training facilitators, e-learning platforms, workshop materials)

- Phase 3: Maturity & Integration (24-36 months)

Activity	Duration	Resource Type
Achieve full technical interoperability across nodes	12 months	People (system integrators, standards coordinators)
Institutionalize governance, metrics, update cycles	12 months	Governance (monitoring leads, platform stewards)
Operationalize translation infrastructure and workflows	12-18 months	People (multilingual translators, QA editors)
Integrate repository with global platforms and policymaking tools	6–12 months	People (policy liaisons, data integration managers)
Launch long-term funding and sustainability strategy	12 months	Governance (funder engagement team, sustainability advisors)

- Strategy 2 : Standardized Record Structures and Interoperable**
- Foundation & Prototyping (0–6 months)

Activity	Duration	Resource Type
Establish technical and advisory groups to oversee standard development	3 months	People (standardization experts, methodologists)
Map existing standards and tools for evidence data sharing	3 months	People (technical reviewers, research assistants); Tools (reference managers, API testing environments)

Develop initial record structure standards for inputs and outputs	6 months	People, Tools (metadata schema builders, validation checklists)
Draft interoperable data formats and exchange protocols	6 months	People (data architects), Tools (format converters)
Prototype tooling for data conversion and structure validation	6 months	People (software developers, statistical programmers); Tools (sandbox environments, validation pipelines)

- 
- Phase 2: Expansion & Feature Enrichment (6-18 months)

Activity	Duration	Resource Type
Build and test retrospective conversion toolkits	6 months	People (developers, QA testers), Tools (extraction and transformation software)
Pilot conversion workflows with selected institutions (e.g., LMIC partners, sectoral repositories)	6 months	People (regional partners, curators), Capacity development (training sessions, implementation guides)

- 
- Phase 3: Maturity & Integration (18-30months)

Activity	Duration	Resource Type
Launch open consultation and finalize standards based on feedback	2 months	People (consultation team), Tools (survey and feedback platforms)
Institutionalize standards through integration into platform governance and training programs	Ongoing	People (WG2, WG5), Capacity development (training modules, onboarding support)

### Strategy 3: Establishing Metadata standards to facilitate data identification and discoverability

#### Foundation & Prototyping (0-6months)

Activity	Duration	Resource Type
Establish metadata working group and regional advisors	3 months	People (metadata experts, multilingual reviewers)
Define metadata schema and classification system	6 months	People, Tools (ontology builders, multilingual term banks)
Map and review existing metadata standards and typologies	3 months	People (metadata reviewers, librarians); Tools (open schema databases, term banks)
Develop strategy for mapping legacy metadata into unified framework	3 months	People (integration specialists, sector experts); Tools (conversion tools, mapping templates)
Draft identifier guidance and linkage protocols	3 months	People (DOI experts, integrators)

#### Phase 2: Expansion & Feature Enrichment (6-18months)

Activity	Duration	Resource Type
Develop tagging tools, including AI-assisted modules	6 months	People (developers, curators), Tools (tagging engines)
Conduct multilingual validation and testing	6 months	People (translators, regional reviewers), Capacity development
Pilot tagging system across selected ESIC-aligned repositories	6–9 months	People (platform partners), Capacity development (training materials)

#### Phase 3: Maturity & Integration (18-30months)

Activity	Duration	Resource Type
Finalize standards through open consultation and feedback	3 months	People (consultation team, regional liaisons)

Institutionalize metadata schema and QA practices across ESIC tools	Ongoing	People (repository managers, WG2/WG5 liaisons)
Maintain multilingual metadata dictionary and update protocols	Ongoing	People (taxonomy curators, QA leads)

#### Strategy 4: Setting Open access standards for equitable data sharing and reuse

Phase 1: Foundation & Prototyping (0-6 months)

Activity	Duration	Resource Type
Convene open access advisory group	3 months	People (legal experts, publishers, LMIC reps)
Draft open licensing and attribution guidance	6 months	People (IP specialists, licensing experts), Tools (template builders)
Conduct legal review of reuse models across jurisdictions	4 months	People (data protection lawyers), Tools (review templates)

Phase 2: Expansion & Feature Enrichment (6-15 months)

Activity	Duration	Resource Type
Build platform capabilities for license display, citation export, and open access tagging	6–9 months	People (developers, UX designers), Tools (repository integrations)
Develop and pilot multilingual access protocols	6 months	People (localization specialists, platform leads), Tools (translation modules)
Host consultations with LMIC institutions and data producers on access priorities	6 months	People (regional facilitators, community orgs), Capacity development

Phase 3: Maturity & Integration (15-24 months)

Activity	Duration	Resource Type
Integrate licensing and attribution norms into platform governance frameworks	Ongoing	People (platform managers, policy advisors)
Monitor and evaluate open access implementation across ESIC systems	Annual	People (independent reviewers), Tools (compliance checklists, usage analytics)
Maintain public guidance materials and licensing templates	Ongoing	People (communications team, legal editors)

#### Strategy 5: Ensuring quality and monitoring and evaluation of data sharing and re-use systems

Phase 1: Foundation & Prototyping (0-9 months)

Activity	Duration	Resource Type
Establish technical QA and M&E advisory group	3 months	People (M&E specialists, equity advisors, platform reps)
Develop draft quality assurance standards	6 months	People, Tools (framework builders, sector-specific guidelines)
Define indicators for system monitoring and ethical adherence	3 months	People (data governance experts), Tools (indicator sets)

Phase 2: Expansion & Feature Enrichment (9-21 months)

Activity	Duration	Resource Type
Build user feedback and issue-reporting interfaces	6 months	People (UX designers, backend devs), Tools (reporting platforms)

Pilot QA and M&E systems in selected repositories	6–9 months	People (QA reviewers, repository leads), Capacity development (training modules)
Conduct participatory review sessions with LMIC and sectoral partners	6 months	People (facilitators, regional advisors), Capacity development

#### Phase 3: Maturity & Integration (21-36 months)

Activity	Duration	Resource Type
Integrate QA/M&E protocols into platform governance and user workflows	Ongoing	People (platform managers, governance leads)
Maintain and update monitoring dashboards and QA tools	Ongoing	People (data analysts, QA leads), Tools (dashboards, repositories)
Conduct periodic system-wide quality and equity evaluations	Annual	People (independent reviewers), Tools (survey and review templates)

### Sustaining the Infrastructure: Funding Models and Incentives

#### Phase 1: Foundation & Prototyping (0-12 months)

Activity	Duration	Resource Type
Map existing funding flows and gaps across ESIC-aligned infrastructure	3 months	People (funding analysts, platform leads)
Convene funders to design pooled or coordinated funding models	3 months	People (funder representatives, facilitators)
Draft incentive framework and potential recognition models	4 months	People (research assessment experts, community leaders)

#### Phase 2: Expansion & Feature Enrichment (12-24 months)

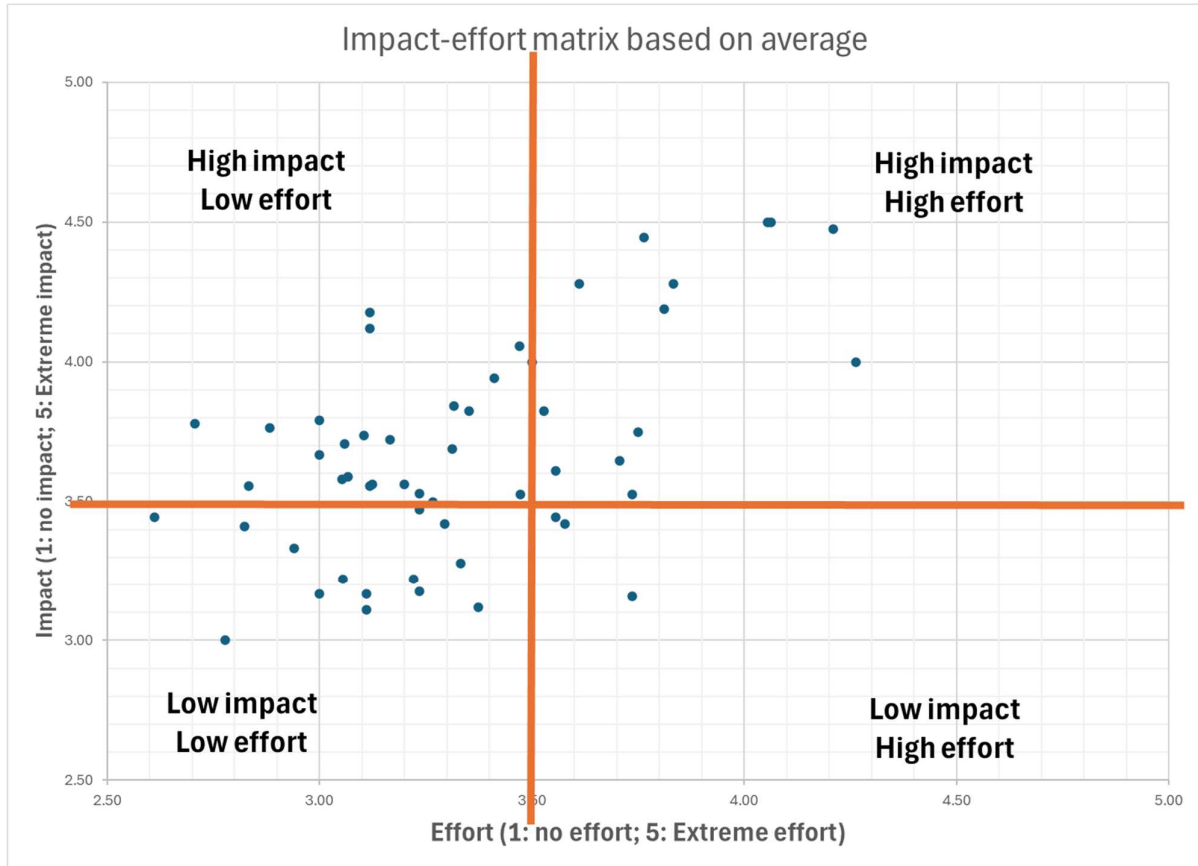
Activity	Duration	Resource Type
Pilot shared grant mechanisms across multiple funders	6–12 months	People (grants officers, financial administrators)
Develop costing templates and sustainability plans for each platform	6 months	People (platform managers, finance teams), Tools (cost modeling templates)
Launch incentive pilots (e.g., citation tracking, recognition in grant calls)	6 months	People (academic institutions, funders), Tools (dashboard features)
Co-design behavioural incentive pilots (e.g., recognition for data curation, open tagging, or use of core outcome sets) with institutions and funders	6 months	People (behavioral science advisors, funders, researchers), Capacity development (community engagement, feedback loops)

#### Phase 3: Maturity & Integration (24-36 months)

Activity	Duration	Resource Type
Institutionalize funding and incentive models across ESIC platforms	Ongoing	People (governance leads, funder liaisons)
Monitor and evaluate sustainability and equity outcomes	Annual	People (independent reviewers), Tools (metrics dashboards, review templates)
Adapt and scale up funding and incentive models based on feedback	Ongoing	People (funder coalitions, platform stewards)



### Appendix 3. Impact-effort matrix



**Appendix 4.** Connection of strategies included in this report with the strategies included in report #3

Strategy	Strategies included from report #3
<b>1. Building a federated repository of living evidence data</b>	3.2.5 Deploy AI-enabled infrastructure to support tagging, content structuring, and extraction
	3.2.1 Federated repository of living evidence data
	3.2.2. Single living repository of data
	3.3.6 User-Friendly Tools for Non-Technical Data Reuse
	3.2.9. Creating comprehensive registries of protocols (of primary studies and evidence syntheses)
	3.2.4 Open API for data integration
	3.3.1 Multilingual user experience and access
<b>2. Developing standardized record structures and interoperable formats</b>	3.6.4 Strategic Funding for Foundational Dataset Production
	3.1.8 Develop standardized record structures for evidence synthesis inputs and outputs
	3.1.5. Evidence synthesis reporting standards
	3.1.6 Evaluation reporting standards
	3.1.7. Primary study reporting standards
	3.1.2 Interoperable data formats and exchange protocols
<b>Establishing metadata standards for discovery and reuse</b>	3.1.9 Retrospective conversion of data and metadata
	3.1.3 Global identifier framework
	3.1.1 Modular metadata standards (to facilitate data curation to be reused)
<b>4. Setting open access standards for equitable data sharing</b>	3.1.4 Metadata transformation tools to support taxonomy alignment.
	3.3.4 Encouraging open licenses and publisher incentives for access
	3.3.5 Establishing organizational open access policies
	3.1.10 Open Access Standards for Data Sharing
	3.2.8 Supporting Open Access Infrastructure
	3.4.2 Encouraging open access through collaborative partnerships to share infrastructure and standard alignment
	3.7.4 Resources Supporting Organizational Shifts to Open Sharing (e.g., toolkits, pilots projects that show the benefits of open sharing)
<b>5. Ensuring quality and monitoring of data sharing systems</b>	3.4.5 Advocacy for Text and Data Mining (TDM) Rights regulation
	3.4.7 Monitoring and Evaluation of Data Sharing System Performance
	3.5.1 Tiered validation processes (to ensure quality and reliability of ES data)
<b>Sustaining the Infrastructure Funding models and incentives</b>	3.5.2 Methodological teams ensuring quality and replicability
	3.6.1 Multi-donor pooled funds to support ES infrastructure
	3.6.5 Sustainable funding for maintaining and updating living data infrastructure
	3.6.3 Recognition mechanisms to incentivize infrastructure contributions
	3.6.7 Link grant eligibility to data standards adherence
	3.6.9 Linked funding models for evidence lifecycle (incl funding for primary research)
	3.6.6 Equitable cost-sharing arrangements (incl models where different partners contribute in different forms)
	3.6.8 Funder and publisher incentives for data sharing
	3.4.6 Align Funding Models with Open Access Principles



**Appendix 5.** Strategies that were finally not prioritized as part of this report (note that some strategies were not necessarily excluded but reframed as part of other strategies included in this report)

- 3.1.11 Streamlined submission standards (single input, multiple uses)
- 3.1.5 Evidence synthesis reporting standards
- 3.2.3. Diagnostic pilots for data integration barriers
- 3.2.6 Provenance tracking systems
- 3.2.7 Technical deployment toolkits for platforms seeking integration.
- 3.3.2 Support offline access and low-bandwidth deployment
- 3.3.3 Regional mentorship models for co-development of platform strategies and synthesis infrastructure governance.
- 3.3.8. Creating translated version of ES outputs to facilitate reuses
- 3.4.1 Data-sharing and reusing agreement templates to build mutual trust
- 3.4.2 Collaborative partnerships for shared infrastructure governance and standard alignment
- 3.4.3 Incentives for conducting QA before data is shared
- 3.4.4 Align reuse systems with ethical and data sovereignty frameworks
- 3.5.3 Crowdsourced and peer-led quality assurance
- 3.5.4 Network of trust for institutional quality signaling
- 3.5.5 Re-use tracking tools using persistent IDs and contributor metadata
- 3.6.2 Coordinated funder platforms for infrastructure alignment
- 3.7.1 Regionally adapted training modules
- 3.7.2 Communities of practice for continuous learning and methodological alignment
- 3.7.3 Pilot demonstrations of reuse value
- 3.7.5 Localized data infrastructure hubs (that provide infrastructure support in LMI settings)
- 3.5.3 Crowdsourced and peer-led quality assurance
- 3.5.4 Network of trust for institutional quality signaling
- 3.5.5 Re-use tracking tools using persistent IDs and contributor metadata
- 3.7.2 Communities of practice for continuous learning and methodological alignment